United States
Naval Postgraduate School

THESIS

ADAPTATION AND VALIDATION OF A COMPUTER RANKING FOR PARTIALLY ORDERED DATA

by

Richard Woodie Mister

Thesis Advisor: J. K. Arima

September 1971

Approved for public release; distribution unlimited.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

ADAPTATION AND VALIDATION OF A COMPUTER RANKING FOR PARTIALLY ORDERED DATA

4. DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Master's Thesis; September 1971

5. AUTHOR(S) (First name, middle initial, last name)

Richard Woodie Mister
Lieutenant Commander, United States Navy

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1971	50	14

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

In 1957 L. R. Ford, Jr., developed a procedure that would produce a rank-order of objects from subjective judgments. Standard procedures usually require that the number of comparisons between any given pair of objects be equal to the number between any other pair. This method does not require any specific number of comparisons between pairs, and it allows that there be missing data. A computer program was developed utilizing Ford's technique. This study adapted the program for use on the IBM 360/67 and evaluated the validity of the program and model which appeared good. Applications for the use of such a program in the Navy were cited.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6811

Unclassified
Security Classification

A-31408

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
rank-order						
missing data						
Computer Program						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0101-807-6821

Unclassified

Security Classification

A-31409

Adaptation and Validation of a Computer Ranking for Partially Ordered Data

by

Richard Woodie Mister
Lieutenant Commander, United States Navy
B.S., United States Naval Academy, 1963

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1971

Author: Richard Woodie Mister

Approved by: James K. Arima, Thesis Advisor

Chairman, Department of Operations Research and Administrative Sciences

Milton U. Clauser, Academic Dean

ABSTRACT

In 1957 L. R. Ford, Jr., developed a procedure that would produce a rank-order of objects from subjective judgments. Standard procedures usually require that the number of comparisons between any given pair of objects be equal to the number between any other pair. This method does not require any specific number of comparisons between pairs, and it allows that there be missing data. A computer program was developed utilizing Ford's technique. This study adapted the program for use on the IBM 360/67 and evaluated the validity of the program and model which appeared good. Applications for the use of such a program in the Navy were cited.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

I. INTRODUCTION

Given, a number of objects to be considered according to the different degrees in which they exhibit some common quality. If the quality is measurable in some objective way, the problem is amenable to treatment by well understood methods. It may happen, however, either for theoretical or practical reasons, that the quality is not measurable, or there is little intuitive feeling as to what form the distribution of the measurements in the population is likely to take. It is then necessary to rely on judgments of a more or less subjective nature carried out after a comparison of the objects among themselves. One method of comparison which has been widely used is that of ranking.

Bradley (1953) purported that criticism of ranking methods stems from a supposed loss of efficiency. When quantitative judgments can be obtained, the magnitude of differences is obscured by the use of ranks. On the other hand, when treatment differences are small and difficult to detect, it would appear reasonable to simplify the procedure for the judge and use a ranking technique. The rank order method is usually computationally simple and often preferred on this ground alone.

Kendall and Smith (1940) investigated a method of preferences where n objects are paired $\binom{n}{2}$ and an observer indicates preference of one object over another. They measured

reliability of judgments on the part of the observer and concordance of preferences between observers. In this treatment they excluded ties. Another comparison technique stemmed from research for the Army demobilization point system. This study by Guttman (1946) covered not only ordinary comparisons but situation comparisons which combined several variates. The developments excluded judgments of equality and assumed that all people compared all pairs. White (1952) presented methods and developed tables for determining the significance of the difference between two treatments in a ranking procedure. This procedure required quantitative values which are ranked and then summed.

One method which has received considerable attention is the method of "paired comparisons." Bradley and Terry (1952) have developed the method of paired comparison for the rank analysis of incomplete block designs. The procedures are applicable where qualitative measurements are reliable and useful in problems involving subjective ranking by judges. No provisions were made for ties or for not ranking a particular pair or group of treatments.

A solution of the ranking problem from binary comparisons developed by Ford (1957) closely paralleled the development by Bradley and Terry. This procedure is singularly important in that it handles problem areas not provided for in any preceding development. Standard procedures usually require that the number of comparisons between any given pair be equal to the number between any other pair. Ford's method does not

require any specific number of comparisons between pairs, and it allows that there be missing data. These two provisions permit considerable flexibility among judges making difficult comparisons.

Ford assumed a matrix $A=(a_{ij})$, where a_{ij} represented the number of times object i had been preferred to object j. Ford associated with each object a weight w_i. These weights would be interpreted as odds, in the sense that the probability of i being preferred to j in a future comparison would be taken to be $w_i/(w_i+w_j)$. With these probabilities, one could compute the a priori probability of obtaining precisely the matrix of results obtained, that is, the matrix A.

In order to determine the set of weights which maximized the likelihood of obtaining matrix A, Ford solved, by an iterative technique, the following equation for each object until the weight stabilized.

$$w_i^{n+1} = \frac{\sum_j a_{ij}}{\sum_j \frac{a_{ij}+a_{ji}}{w_i^n+w_j^n}}$$

where a_{ij}= number of times object i was preferred to object j; a_{ji}= number of times object j was preferred to object i; and, w_i^n= weight assigned to object i on the n^{th} iteration. Ford made the following assumption about matrix A: "In every possible partition of the objects into two nonempty subsets, some objects in the second set has been preferred at least once to some object in the first set." In order to yield a solution the data must meet this criterion.

In the early 1960's in an effort to determine a rank-ordering of measures of scientific performance, Pelz and Andrews (1966) developed a computer program which embodied the Ford procedure. In order to satisfy the partitioning assumption by Ford, the program incorporated a means for separating universal highs and universal lows prior to computation of weights. The other ways in which the assumption could be violated were if some objects were not rated relative to other objects or if some objects would fall in a subset such that comparisons were all in one direction. Addition of a small constant to each cell of the matrix A solved the last two violations. The computer program developed did not provide a means to maintain the identity of each judge nor examine the consistency of the judges with one another.

The purpose of this study was three-fold:

1. Adaptation of the present computer program for use on the IBM 360/67 computer.
2. Statistical validation of the program and model.
3. To indicate the implications of adapting such a program for use in the Navy.

The assumption was made a priori that the capability of judges was uniform throughout the experiment.

II. METHOD

A. TEST CATEGORIES

Proper validation of the program required testing and a comparison of the program results with a known true order of items or a universally accepted standard. Four test categories of verbal items were selected in which the items listed were highly familiar to all subjects tested. These categories contained items which had at least a .9 correlation over test subjects in the category norms for verbal items compiled by Battig and Montague (1969).

In order to insure that there would be ties and missing data in the testing, the following criterion was used for item selection within each category. First, items were grouped together into approximately four groups. The criterion between items in each group was a one to three percent change in frequency of occurrence based on the Thorndike and Lorge (1944) general count. The second criterion was a five to ten percent change in the frequency of occurrence between groupings. These criteria yielded essentially a type of clustering in four ranges of frequency of occurrence of verbal items.

To compare things two at a time and judge which has higher rank or to rank all n things simultaneously, that is judging n(n-1)/2 comparisons at once, are substantially equivalent procedures. However, comparing two things at a time

allows inconsistencies (intransitivity) to appear within judgments of individuals, and it is sometimes harder in practice for people to judge n things simultaneously than to compare them two at a time. In order to eliminate these two problems, items were not presented in pairs, Rather, the entire category was included on one testing sheet whereby each individual could see all items, essentially combining the two methods. Appendix B shows the categories and items in each category in order of their frequency of occurrence.

B. SUBJECTS

Twenty male and female subjects ranging in age from 24 to 37 years, with comparable levels of education, were selected. Each subject was used twice. Ten subjects were assigned at random to each of four test categories.

C. PROCEDURE

Subjects were given a standard set of instructions explaining the ranking procedure to be used (Appendix A). Subjects were required to work rapidly and to give their first impression as to assignment. The items were placed on the test sheet in a random order. The subject's replies were recorded on the testing sheet.

D. DESIGN

Compiled data was input to the program and an overall ranking obtained. The method of evaluations used, when not all items are compared by all judges, indicates a sign test or a signed-rank test should be used [Abelson and Bradley

1954]. The use of Analysis of Variance was not appropriate. In the formation of subjective tests the assumptions of Analysis of Variance are seriously suspect [Bradley 1955]. Analysis of Variance also requires quantifiable data, not qualitative data. Dixon (1953) has shown that the sign test compares more favorably with Analysis of Variance for small samples than indicated by results on relative efficiency. In investigating power of paired comparisons, it was found that the efficiency of the method of paired comparisons relative to Analysis of Variance, and under conditions appropriate to Analysis of Variance, was $t/\pi(t-1)$ where t is the treatment being considered. When $t=2$ the efficiency reduces to $2/\pi$, the relative efficiency of the sign test.

Data results were amenable to a Wilcoxon signed rank test with a hypothesis that the treatments are equal. A Spearman rank correlation was conducted and a hypothesis that rho=o was tested. The Kendall coefficient of concordance was not used but is similar to the Spearman rank order correlation using a hypothesis that tau=o. All tests were done at a significance level of .05.

E. COMPUTER PROGRAM

The computer program performed an overall rank-ordering of partially ordered data. The program is set up in three parts; the main program and two subroutines.

1. Main Program

The main program performed essentially two functions after reading all the input data. Beginning with the first

judge, a sequential ID number was assigned to the original ID number of objects that were judged in order of their appearance. The procedure continued until all objects were accountable. No duplication of assignments were made. Assigned ID numbers were used throughout the program and the original ID numbers stored until the final printout of weights.

Beginning with each judge a count was made of the number of comparisons which were to be made between each pair of objects. No comparisons were made between objects tied, that is, objects assigned to the same rank.

2. Subroutine- CORE2

The first section tabulates for each individual comparison the number of times that comparison was made by all judges in the experiment. It was done sequentially from the input ranking-order of the judges and was the number of times object i was preferred to object j, that is, the win-loss matrix.

The next section through a series of logic switches determined which objects were rated "universal highs" or "universal lows" and removed them from the weight calculation. The appropriate rows and columns of the win-loss matrix were also removed.

The next section computed the initial weighting factor from the win-loss matrix. To each cell in the matrix a small constant was added, .00001, to prevent violation of Ford's partition assumption.

3. Subroutine- CORE3

In this subroutine the final calculations were made for the new weighting factor and then the new weighting factor was compared to the old factor to determine if the preset convergence criterion was met. (No objects weighting factor changed more than .005 between iterations.) If this criterion was not met then a count of the number of iterations was made against the number input in order to determine whether the program would terminate without convergence.

III. RESULTS

The experimental results are compiled in Appendix C. These results were studied in an attempt to pinpoint any significant differences between the experimental results and the accepted standard of Thorndike and Lorge. It should be remembered that these results pertain to the specific type of testing used. The results did indicate some obvious similarities between the methods and permit some fairly general conclusions.

Three means are available to analyze ranked data; the Wilcoxon signed rank test, the Spearman rank correlation, and the Kendall Coefficient of Concordance. All three tests measure essentially the same relationships between the sets of data, however, the method is different.

The signed rank test takes into account the magnitude of the observed differences between the data sets. The hypothesis tested at the .05 significance level was that there was no difference between the effects of the two treatments. This hypothesis was accepted for all four categories.

For ease of computations the Spearman rank order correlation was used in lieu of the Kendall coefficient of concordance. However, the same conclusion would be reached, namely to accept or reject the null hypothesis, by computing the Kendall coefficient of concordance [Ostel 1963]. The Spearman rank order correlation yielded the following for each category tested:

Category	rho
I	.521
II	.598
III	.687
IV	.460

The hypothesis tested was rho=o. In Category II and Category III the hypothesis was rejected at the .05 level when a two tailed t test was used. In Category I and Category IV the computed t value indicated the failure to reject a hypothesis of rho=o. This can be explained, however, by looking at the rankings of the program and the Thorndike and Lorge count. In Category I the computed difference in score between only one pair was excessive. This was for the material "felt." In view of the present day use of synthetic materials it is believed that the familiarity with felt material is quite low compared to what it was 26 years ago. In Category IV there were two diseases, typhoid and syphilis, which were exactly reversed in the program ranking. This reversal yielded a large difference in score between each pair. Today's cleanliness standards and medical developments have lessened the familiarity with typhoid which would place it low in a current ranking. Due to the fact that the majority of the test subjects were military personnel with a broad background, a greater familiarity with syphilis and a tendency to place this item high on any ranking list could be expected.

Recomputation of the correlation coefficient for Category I with "felt" removed, yielded a rho= .788. In this case the t statistic indicated rejection of the hypothesis that rho=o.

Recomputation of the correlation coefficient for Category IV with either "typhoid" or "syphilis" removed or reducing the difference in score between each of the two pairs to one-half of its present value yielded a rho= .585. The t statistic indicated rejection of the hypothesis that rho=o.

Rank-order stability was reached after the first iteration for Category I, III, and IV. Stability was reached for Category II after the third iteration. Category I converged in thirty-five iterations and Category III converged in sixteen iterations. No convergence was reached for Category II and Category IV after fifty iterations. Figures 1, 2, 3, and 4 show how weights of each object changed over fifty iterations for each category. Four objects in Category III were rated as universal highs and were removed prior to computation of weights.

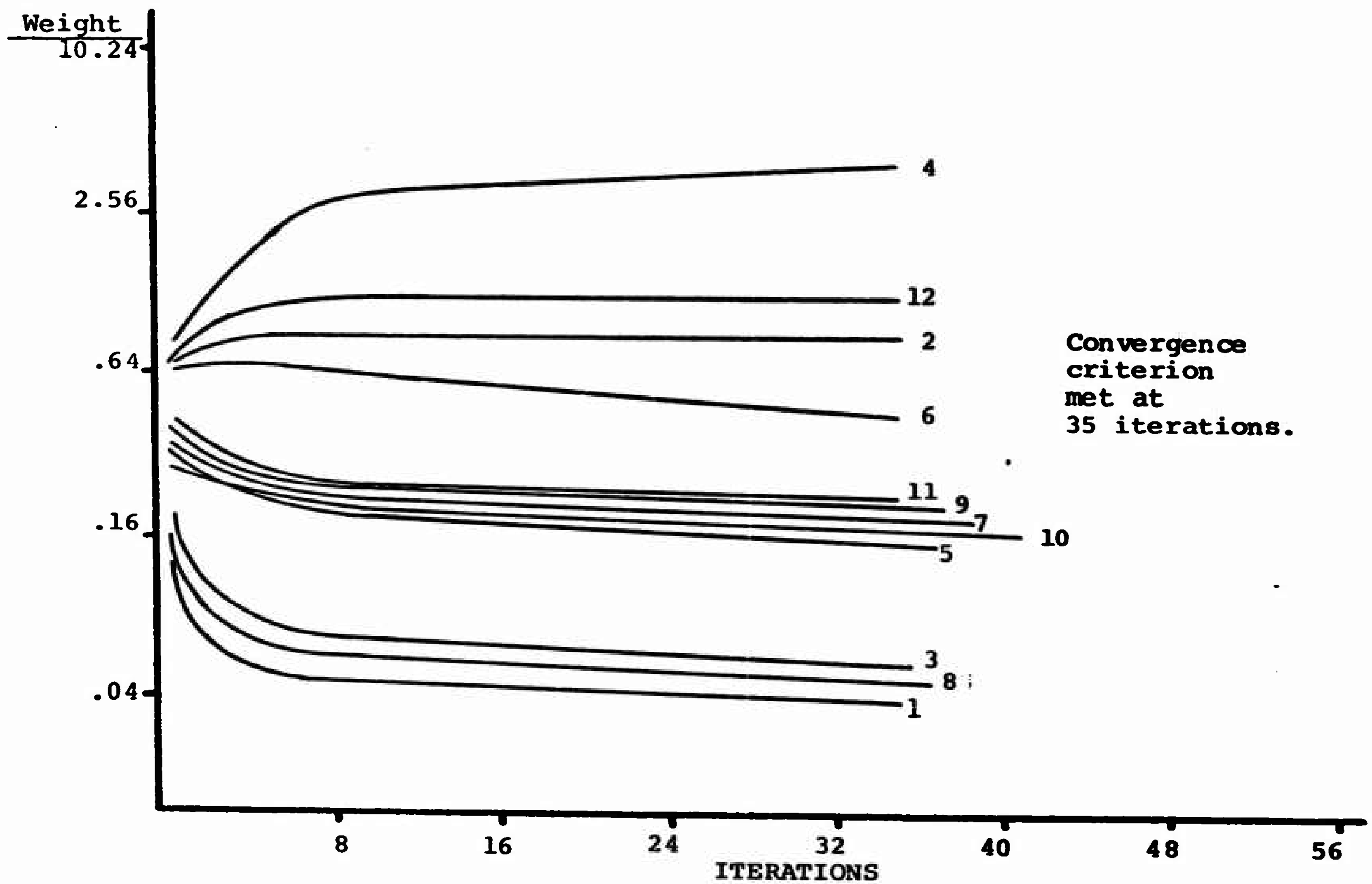

Figure 1. Weight Change per Iteration. Category I. Assigned ID Number Used.

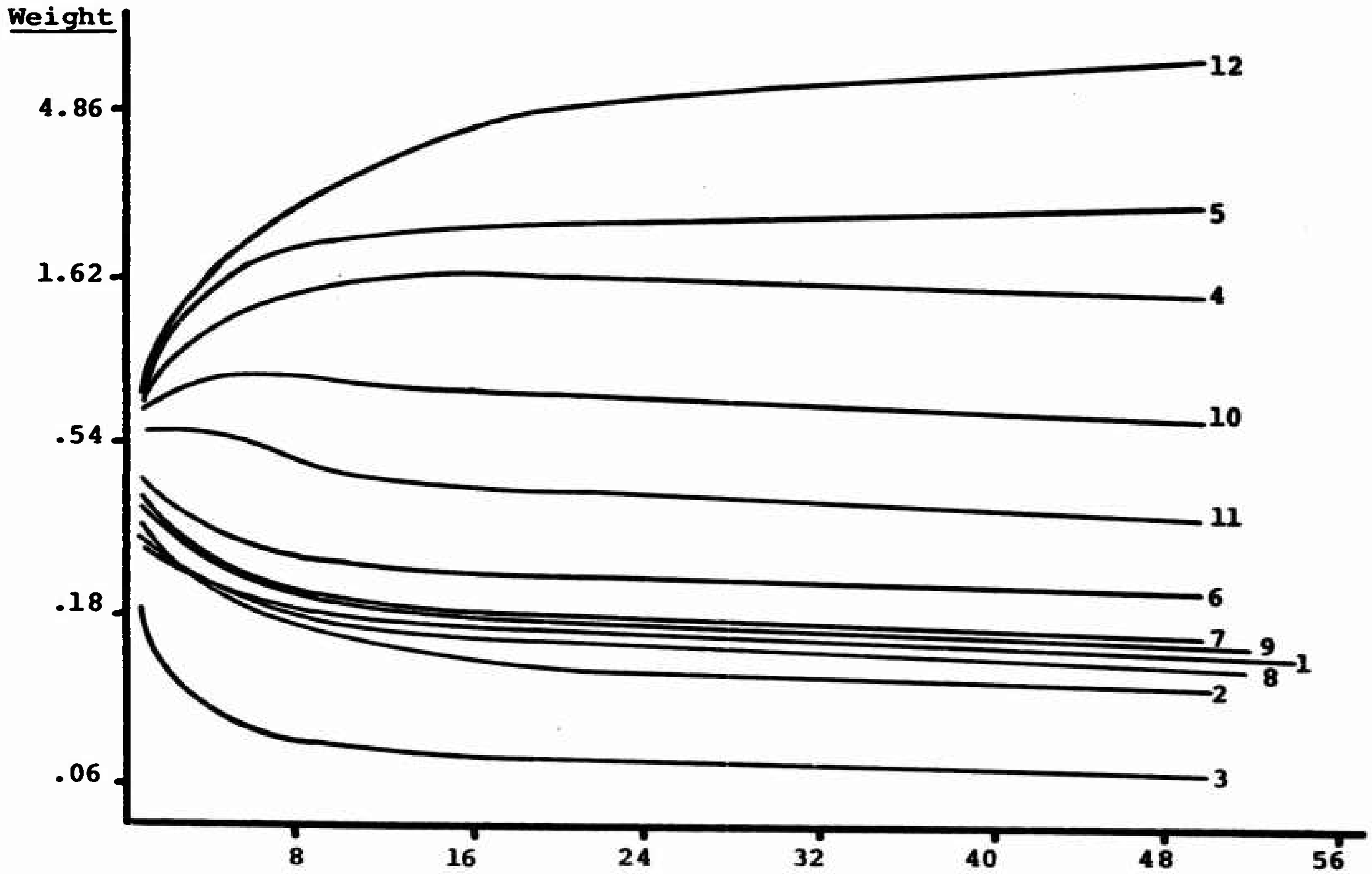

Figure 2. Weight Change per Iteration. Category II.
Assigned ID Number Used.

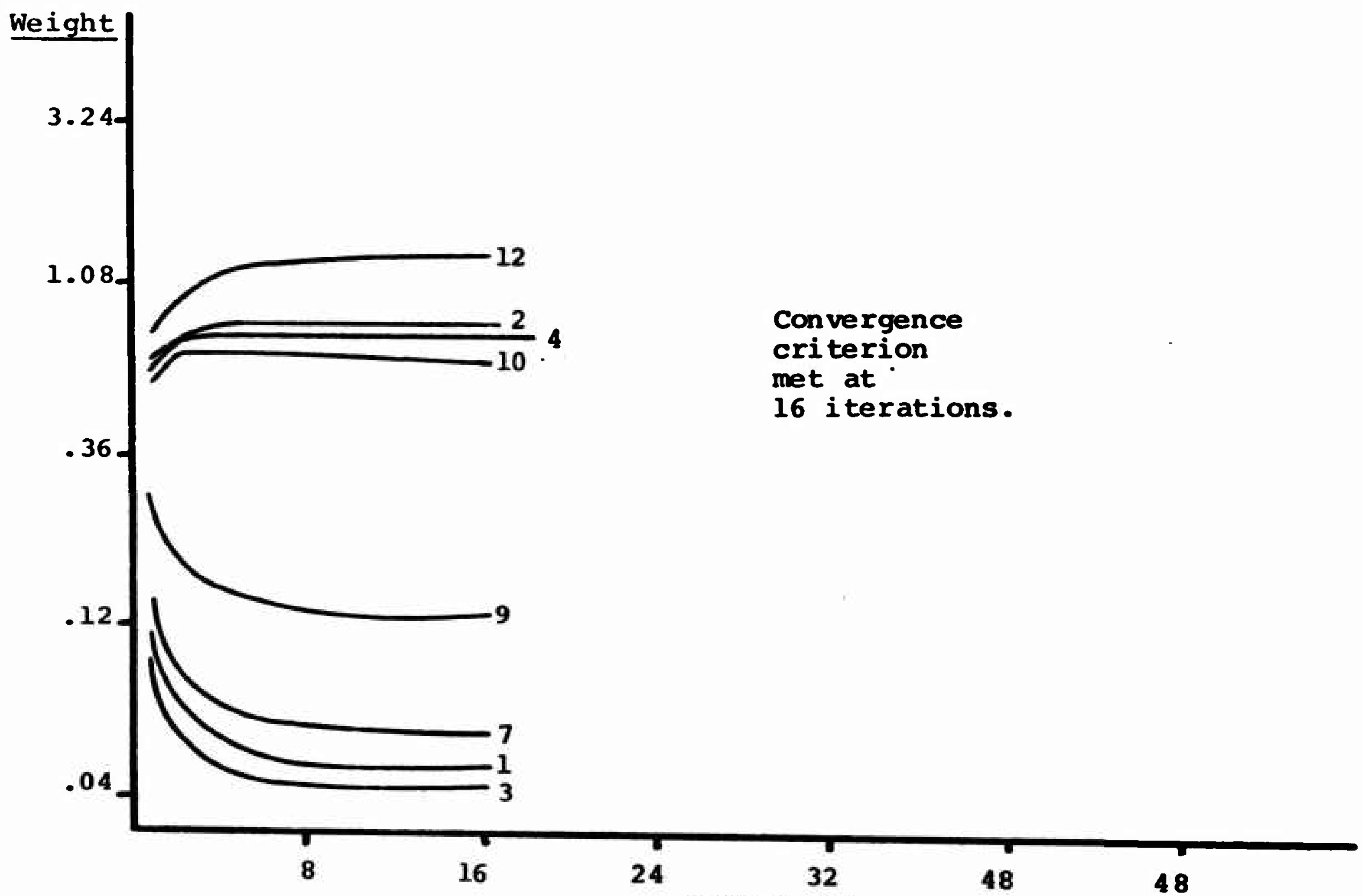

Figure 3. Weight Change per Iteration. Category III. Assigned ID Number Used.

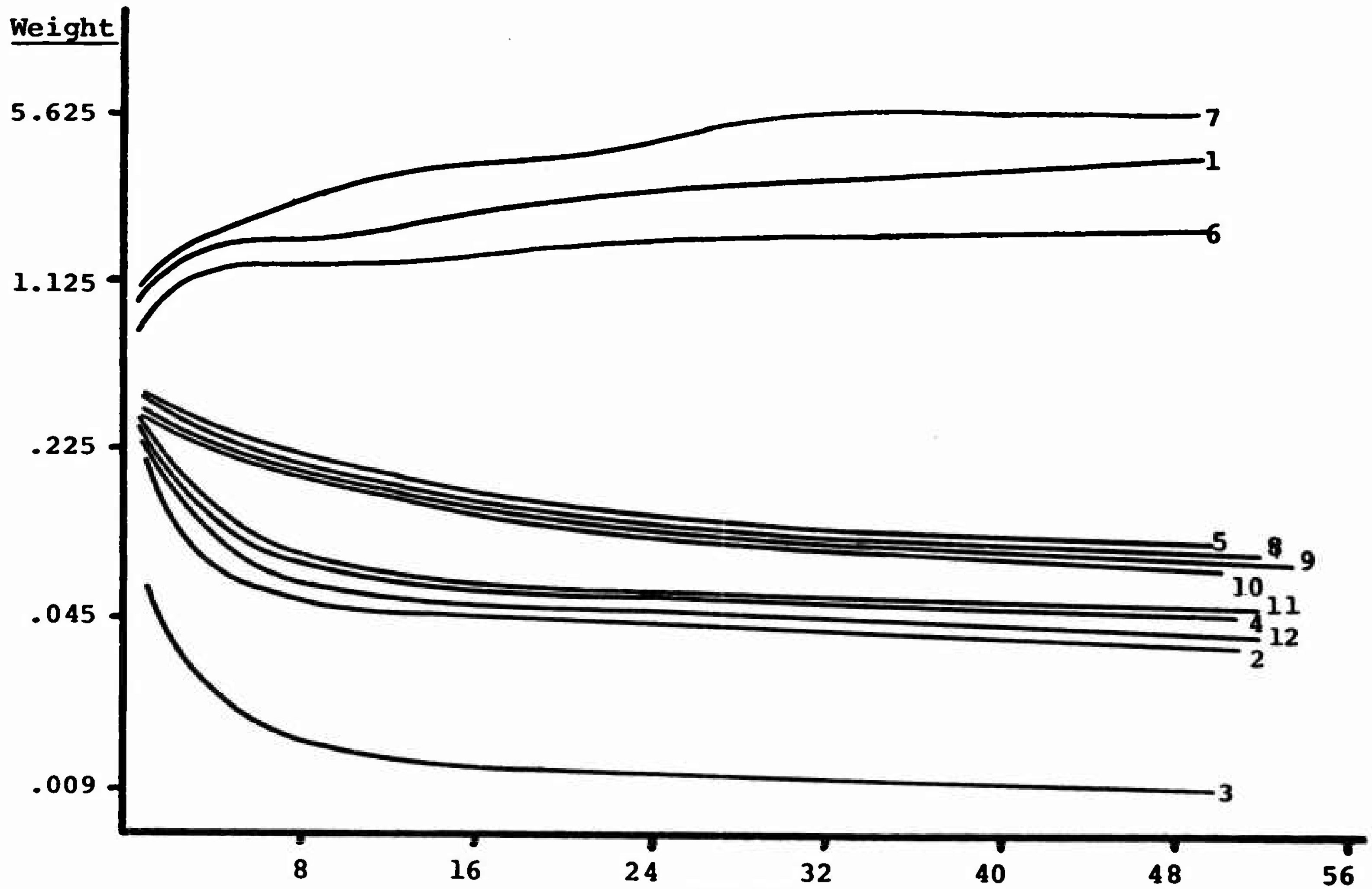

Figure 4. Weight Change per Iteration. Category IV. Assigned ID Number Used.

IV. DISCUSSION

Evaluation of the results by both the Wilcoxon signed rank test and the Spearman rank order correlation indicated that the program produced the proper rank order for the items tested. The results obtained were statistically valid and appear consistent with real world data. The results can only be considered approximately correct for the categories examined, since they were subject to variations in the ranking by judges. One source of variation occurs when the quality being measured is not known with certainty to be represented as a linear variable. An observer may rank a number of objects on this quality believing that he is doing something within his powers. However, if this quality is not measurable on a linear scale, the ranking may fail to give a real picture either of the observer's preference or of the variation of the quality among the objects. Another source of variation is when an observer produces a configuration of preferences which show inconsistencies. There are usually several explanations; he may be an incompetent judge, the objects may be so alike that consistent differentiation is not possible, or his attention may wander during the course of the experiment. The second source of variation was not considered in the experiment since it was assumed a priori that the standards of judging were uniform throughout the subjects being tested. The first means of variation was not a factor due to the way in which the items were selected.

The correlations obtained were not expected to be very high due to the method of selecting items to insure ties and missing data. This procedure for item selection provided a check of the program's method of ranking. Another reason for somewhat lower correlations was that the reference material used for comparison [Thorndike and Lorge 1944], is somewhat outdated.

V. APPLICATIONS FOR USE OF PROGRAM IN U.S. NAVY

There are many instances in the Navy where the usefulness of such a program may prove invaluable as a labor saving device. Ranking procedures are used throughout the Navy in various forms. Although none of the systems in use have been examined in order to determine their relative efficiency, each system requires a large amount of manhours, and the results may still be biased by many factors unknown to the individual or agency assembling the overall ranking.

A. SPECIFIC USES

Annually the Navy has proposed to it or makes proposals for various research programs. The amount of money spent and the number of feasibility studies undertaken in order to determine which programs should have priority for development and which ones should be discarded is not known. It is easy to imagine how a ranking program might be used to determine which proposals should be put into committees for further study and evaluation.

Many times each year military officers are available for assignment to new billets and changes of duty stations. In order to determine proper assignment, a ranking procedure is used taking into account an officer's performance based on fitness reports, his desires from a preference card, and several other factors. The computer program cited would facilitate a large reduction in manhours spent tabulating this

data and ranking those officers within the group for assignment. A modification to the existing program would be required in order to give more weight to certain factors and to provide a means of weighting the competence of certain judges.

When military personnel are transferred there are many questionnaires which are filled out rating the supply facility which handled the movement of their household goods and the shipping firm which did the actual moving. A ranking procedure in this case would point out which facilities are doing a good job and which ones are substandard.

The preceding paragraphs have pointed up three of many uses for which a computer ranking procedure may be used.

B. AN EXAMPLE

The Naval Personnel and Training Research Laboratory at San Diego and the Naval Personnel Research and Development Laboratory at Washington, D.C. annually publish documents which describe programs which have been developed or have been contracted for by the Navy. From these documents 10 titles were randomly selected and abstracts prepared describing the programs selected. Ten Naval officers from the Naval Postgraduate School were asked to rank these programs on their desirability and need for retention and further development within the Navy. Subjects were given a standard set of instructions to be used in the ranking procedure (Appendix D).

VI. CONCLUSIONS

The computer program cited provided a simple and easy means for combining sets of partially ordered data. The program produced a rank-order which appeared consistent with the real world order of the objects that were ranked. Tests indicated that there was no significant statistical difference between the rank-order produced by the program and the true order. This result must remain somewhat tentative in view of the fact that more extensive experimentation was not conducted. The assumption that there were uniform standards of judging throughout the subjects being tested was logical in view of the test results.

A ranking program of this nature would be valuable for use in the military. It could be used to replace or supplement existing ranking procedures which are now used. The simplicity of such a program would yield reductions in manhours and costs of most systems now in operation. In some situations modifications to the program would be required to include a provision for giving additional weight to certain factors which would be more important in the ranking procedure.

APPENDIX A

INSTRUCTIONS TO JUDGES-TESTING PHASE

The purpose of this experiment is to achieve an ordinal ranking of similar objects which belong to different categories. Each category will contain a list of twelve objects belonging to that category. For each category, you are to make an ordinal ranking of the objects in that category as to what you believe their relative familiarity is to all people in general. You are requested to judge only those objects which you can rank with confidence. You are permitted to use as many ordinal ranks for each category as you deem necessary, and to place as many objects in each rank as you choose. In order to simplify the procedure, after looking at the words in each category, select the number of ordinal ranks which you will use. Write the number of the rank next to the object you are assigning to that rank for the objects you choose to judge. Work as rapidly as possible and give your first impression as to assignment.

Are there any questions?

APPENDIX B

ACCEPTED STANDARD RANK ORDERING OF ITEMS BY CATEGORY

CATEGORY I

(4)	cotton
(5)	felt
(12)	wool
(2)	lace
(9)	velvet
(6)	canvas
(8)	muslin
(3)	pique'
(10)	rayon
(11)	corduroy
(7)	denim
(1)	batiste

CATEGORY II

(10)	cup
(11)	bowl
(4)	knife
(12)	fork
(9)	refrigerator
(5)	saucer
(3)	sieve
(6)	skillet
(8)	ladle
(2)	scraper
(7)	toaster
(1)	cleaver

CATEGORY III

(8)	salt
(6)	sugar
(9)	sage
(10)	ginger
(11)	vinegar
(12)	cloves
(5)	mustard
(2)	cinnamon
(4)	nutmeg
(3)	thyme
(7)	basil
(1)	cayenne

CATEGORY IV

(7)	cold
(9)	rheumatism
(4)	typhoid
(1)	cancer
(10)	smallpox
(11)	cholera
(6)	measles
(2)	rheumatic fever
(5)	syphilis
(8)	diabetes
(12)	dysentery
(3)	peritonitis

ID numbers assigned to items in each category are in parenthesis.

APPENDIX C

COMPUTER RANK ORDERING
OF ITEMS BY CATEGORY

CATEGORY I		CATEGORY II	
(4)	cotton	(12)	fork
(12)	wool	(5)	saucer
(2)	lace	(4)	knife
(6)	canvas	(10)	cup
(11)	corduroy	(11)	bowl
(9)	velvet	(6)	skillet
(7)	denim	(7)	toaster
(10)	rayon	(9)	refrigerator
(5)	felt	(1)	cleaver
(3)	pique'	(8)	ladle
(8)	muslin	(2)	scraper
(1)	batiste	(3)	sieve

CATEGORY III		CATEGORY IV	
(8)	salt	(7)	cold
(6)	sugar	(1)	cancer
(5)	mustard	(6)	measles
(11)	vinegar	(5)	syphilis
(12)	cloves	(8)	diabetes
(2)	cinnamon	(9)	rheumatism
(4)	nutmeg	(10)	smallpox
(10)	ginger	(11)	cholera
(9)	sage	(4)	typhoid
(7)	basil	(12)	dysentery
(1)	cayenne	(2)	rheumatic fever
(3)	thyme	(3)	peritonitis

ID numbers assigned to items in each category are in parenthesis.

APPENDIX D

INSTRUCTIONS TO JUDGES
APPLICATION EXAMPLE

The purpose of this experiment is to achieve an ordinal ranking of the desirability of development, maintaining, and/or continuing certain research programs within the U.S. Navy. You will be given a description of ten research programs currently in use or being proposed by the various research laboratories in the Navy. You are to make an ordinal ranking of the programs as to what you believe their desirability and need are for retention and further development within the Navy. You are requested to judge those areas with which you feel you can rank with confidence. You are permitted to use as many ranking categories as you deem necessary, and to place as many programs in each category as you choose. In order to simplify the procedure, after reviewing the programs, select the number of ranking categories which you will use. Write the number of the category next to the program you are assigning to that category for the programs you choose to judge. Work rapidly and give your first impression as to assignment.

Are there any questions?

APPENDIX E

COMPUTER RANK ORDERING OF RESEARCH PROJECTS
APPLICATION EXAMPLE

105 TITLE: Improved Enlisted Personnel Distribution and Management.

DESCRIPTION: A computer assisted distribution and assignment (CADA) system is being designed to help improve the utilization of enlisted manpower. Preliminary model currently is being implemented in the Pacific Fleet. Prototype model is now under development for application in BUPERS in support of centralized management of enlisted ratings. Related research results include development of computer and mathematically based procedures for (1) the equitable allocation of personnel resources, (2) the optimal match of man and billet, (3) the identification of billet vacancies in order of priority, (4) the projection of the number of distributable assets, and (5) the feedback of information on the results of distribution management actions.

101 TITLE: Ship Manning Requirements Techniques

DESCRIPTION: The increasing sophistification and complexity of naval ships, systems, and equipments in the face of project volunteer and a smaller Navy requires the development of methods which will improve the accuracy of manpower requirements forecasting and manpower utilization.

A technique for defining and documenting manpower requirements for ships based on the application of selected work study techniques to basic manning criteria in each of the separate work areas aboard ship has been developed. It permits the production of a document which displays in detail the rationale for manning by ship classes based on equipment and required operational capabilities to meet mission assignment.

104 TITLE: Evaluation of Standards for Navy Reenlistment.

DESCRIPTION: This research was generated out of concern over the quality of reenlistees. Unsatisfactory performance was costing the military services enormous amounts of money in such things as reenlistment bonuses and pay and allowances for reenlistees from whom commensurate service was not realized. Court and confinement costs of reenlistees were cited. It was suspected that personnel of inferior quality were being allowed to reenlist, including some with unsatisfactory first term records.

In an attempt to identify unsatisfactory individuals prior to reenlistment, comparisons were made between unsatisfactory and satisfactory reenlistees on information available at the time of the reenlistment decision. The project also provided information on the effect on manning which would result if reenlistment standards were made more stringent.

102 TITLE: Development of Navy Military Personnel Costing Techniques for Use in Determining Cost Implications Associated with Changes in Reenlistment Rates.

DESCRIPTION: Thousands of skilled technicians are required to operate and maintain the complex systems and equipment now in the Fleet. The Navy constantly experiences difficulty in retaining these technicians because of competition for them from other sectors of the economy.

To alleviate this problem, several technician-oriented procurement programs and career incentive programs are employed. To facilitate evaluation of these programs, a methodology for determining the relative cost benefits associated with retention of personnel has been developed.

103 TITLE: Design of an Optimum Personnel Force Structure.

DESCRIPTION: An optimum force structure containing appropriately qualified personnel in sufficient numbers at least cost cannot now be certified. This project is concerned with the development of improved techniques to analyze and balance the relationship between personnel requirements and the composition of the existing force structure.

106 TITLE: Interest Measurement in Officer Selection.

DESCRIPTION: Each year several thousand young men apply for officer training programs at the Naval Academy and NROTC units at various colleges. High attrition rates

are experienced in both training and active duty. To reduce the cost of losing substantial proportions of these men, it is imperative that those applicants having the greatest career potential be identified in the selection process. Several years of research on vocational interest tests and biographical questionnaires have made it possible to identify those applicants most likely to successfully complete officer training and remain in the Navy after completing their minimum requirements.

110 TITLE: Evaluation Survey of the Effectiveness of Submarine Sonar Operator Training.

DESCRIPTION: A comprehensive survey was accomplished of the proficiency, training, and utilization of submarine sonar technicians and sonar watchstanders. The survey provided up-to-date information concerning the efficiency of training procedures. Such information is necessary on a periodic basis to insure appropriate alignment of the training to fleet requirements in order to prevent serious impairment of operational fleet submarineASW efficiency. Data gathering instruments included interview forms, self ratings, supervisor ratings, knowledge tests, and performance tests.

107 TITLE: Marginal Personnel/Minority Group Testing.

DESCRIPTION: Present test batteries used in both military and civilian settings have been criticized for alleged inequities when used with groups defined on the basis of race or ethnic affiliation. Public policy as well as

efficient manpower utilization requires that all personnel be afforded equality of opportunity in assignment and that those abilities being measured bear relevance to skills required on-the-job.

109 TITLE: Personnel Cost Research for Early Man/Machine Design Trade-Offs.

DESCRIPTION: The critical element of personnel cost has not been systematically considered when making system design and development decisions early in the system development cycle. No tools exist to enable the cost-effectiveness of such decisions to be measured. For this reason, research was undertaken to develop a personnel cost model for use in personnel and man-equipment trade off decisions. A basis model was accomplished which allowed the identification of all pertinent cost items and the accumulation of cost elements in an unequivocal manner.

108 TITLE: LOFARGRAM Analysis Procedures.

DESCRIPTION: The airborn JEZEBEL system has shown great potential as a means of detecting and classifying underwater contacts; however, its usefulness has been continually hampered by the lack of adequately trained operators. One of the main reasons for operator deficiencies is that training programs have been seriously hampered by the lack of a standardized, systemic procedure for analyzing the information displayed on the gram which is the main display component of the system.

In order to correct this situation, a systematic LOFARGRAM procedure was developed.

ID numbers assigned to abstracts are to the left of the title.

APPENDIX F

EVALUATIONS OF ITEMS PER CATEGORY BY TEN JUDGES

CATEGORY I

True Order	Object	Evaluations by Judges: I	II	III	IV	V	VI	VII	VIII	IX	X
1	4	1	1	1	1	1	1	1	3	2	1
2	5	4	3	1	2	7	2	4	7	4	2
3	12	2	2	1	1	2	1	1	1	5	1
4	2	5	2	1	2	4	2	4	2	1	1
5	9	6	3	1	2	5	1	3	9	4	1
6	6	3	2	1	3	3	1	2	5	5	1
7	8	10	4	2	5	6	2	5	8	3	3
8	3			3		8	3		11	1	3
9	10	7	3	2	4	5	1	2	10	1	2
10	11	8	3	1	3	4	1	3	6	4	1
11	7	9	2	2	5	5	2	2	4	3	2
12	1			3		8	3		12	2	3

CATEGORY II

True Order	Object	Evaluations by Judges: I	II	III	IV	V	VI	VII	VIII	IX	X
1	10	3	1	1	1	2	1	2	1	2	1
2	11	1	1	2	1	2	1	2	1	2	1
3	4	6	1	1	1	1	1	1	1	1	1
4	12	2	1	1	1	1	1	1	1	4	1
5	9	5	2	2	1	3	1	4	2	5	1
6	5	4	1	1	1	5	1	2	3	4	1
7	3		3	4	6	7	2		3	3	2
8	8	9	4	4	3	7	2	1	4	4	2
9	6	8	2	3	6	6	1	3	4	3	1
10	2	10		4	4	7	1	1	5	3	3
11	7	7	2	2	1	4	1	3	2	5	1
12	1	11	3	3	5	7	3	1	5	5	2

CATEGORY III

Evaluations by Judges

True Order	Object	I	II	III	IV	V	VI	VII	VIII	IX	X
1	8	1	1	1	1	1	1	1	1	1	1
2	6	2	1	1	1	1	1	1	1	1	1
3	9	10	5	3	5	5	5	6	3	3	4
4	10	8	3	4	3	4	3	5	3	2	1
5	11	4	2	2	2	3	2	3	2	1	1
6	12	6	3	3	3	4	3	4	2	2	3
7	5	3	2	2	1	2	2	2	2	1	1
8	2	7	4	3	2	4	2	4	5	1	2
9	4	5	3	3	4	4	4	5	3	2	2
10	3		6	3	5	7	5	7	4	3	4
11	1			5	5		4	7	4	3	4
12	7	9	6	5	5	6	5	7	4	3	4

CATEGORY IV

Evaluations by Judges

True Order	Object	I	II	III	IV	V	VI	VII	VIII	IX	X
1	7	1	1	1	1	1	1	1	1	1	1
2	9	1	1	2	4	5	9	5	4	4	2
3	4	3	2	3	3	8	6	4	4	5	2
4	1	1	1	1	1	3	1	3	2	2	1
5	10	3	2	4	2	10	2	4	4	2	2
6	11	3	2	2	3	11	6	4	4	5	2
7	6	1	1	1	2	2	2	2	3	3	1
8	2	2	2	6	5	4	5	5	4	5	3
9	5	2	1	3	2	7	3	5	4	4	2
10	8	1	1	3	4	6	4	5	4	4	2
11	12	3	1	5	2	9	7	6	5	3	2
12	3	2	2	7	6		3	7	4		3

APPENDIX G

EVALUATION OF ABSTRACTS BY TEN JUDGES

	Evaluations by Judges									
Abstract	I	II	III	IV	V	VI	VII	VIII	IX	X
101	2	1	1	5	2		1	2	1	7
102	1	2	2	4	3		1	2	3	2
103	2	1	1	6	2		3	1	4	4
104	1	1	3	7	3		1	1	1	5
105	2	2	1	1	1	1	2	1	2	1
106	3	3	1	3	2	2	1	1	5	6
107	3	1	2	8	2		4	3	2	10
108	2	2	3		1	1	4	3	4	8
109	2	3	2	2	3		3	2	4	9
110	1	2	3		3	2	4	3	3	3

```
C          A COMPUTER RANKING PROCEDURE FOR PARTIALLY ORDERED
C          DATA. CCMPARISONS MADE IN THREE CORES.
C               INPUT VARIABLE NAMES-
C                N= # OF OBJECTS BEING RANKED   < OR = TO 130
C                JJ= # CF JUDGES   < OR = TO 130
C                EPSLON= CONVERGENCE CRITERION
C                JUPPER= MAX # OF ITERATION> DESIRED W/OUT
C                        CCNVERGENCE   .
C          THE DATA DECK IS ARRANGED IN THE FOLLOWING MANNER-
C             1.LABEL CARD- PUNCH 1 IN COL 1, FOLLOWED BY 71H
C             2.PARAMETER CARD- FORMAT(2I6,F6.3,I6)
C               COL 1-6 = # OF OBJECTS (N)
C               COL 7-12 = # OF JUDGES (JJ)
C               COL 13-18 = CONVERGENCE CRITERION (EPSLON)
C               COL 19-24 = MAX # OF ITERATIONS (JUPPER)
C             3.JUDGE CARD- FORMAT(I6)
C               COL 1-6 = # OF RANKS USED BY THIS JUDGE (NG)
C             4.DATA CARDS- FORMAT(24I3)
C               COL 1-3 = # CF OBJECTS IN THIS RANK
C               COL 4-72 = ID NUMBER OF OBJECTS IN THIS RANK,
C            CONTINUE ON WITH AS MANY CARDS AS NECESSARY TO
C            COMPLETE A RANK.
C             5.THERE MUST BE (JJ) BLOCKS OF (3) AND (4).
C
      INTEGER*2 LIST(130)
      DIMENSICN NR(130),NC(130),NTG(130),MC(130,130)
      CCMMON LIST
      CCMMON NG,NGM1,IP1,NTGI,NTGJ,NTG,EPSLON,NR,NC,N,NCOUNT,JUPPER,MC
      INTEGER CUTPUT
      LEN=0
      INPUT=5
      IDISK=9
      OUTPUT=6
      READ(INPUT,1000)
 1000 FCRMAT (72H
     X                                                              )
      WRITE(OUTPUT,1000)
      REWIND 9
      READ(INPUT,1) N,JJ,EPSLON,JUPPER
    1 FCRMAT (2I6,F6.3,I6)
      DO 2 I=1,N
      NR(I)=0
    2 NC(I)=0
      NCOUNT=0
      DO 8 J1=1,JJ
      NCNT1=NCOUNT
      READ(INPUT,3) NG
    3 FCRMAT (I6)
```

```
C                                                                       C
C                                                                       C
C        A SEQUENTIAL ID NUMBER IS ASSIGNED TO THE ORIGINAL             C
C        ID NUMBER OF THE OBJECTS BEING JUDGED IN ORDER OF              C
C        THEIR APPEARANCE IN THE DATA, UNTIL ALL OBJECTS                C
C        BEING JUDGED ARE ACCOUNTED FOR. NO DUPLICATION OF              C
C        ASSIGNMENTS IS MADE.                                           C
C                                                                       C
      DO 5 NG1=1,NG
      READ(INPUT,4) NTG1,(MC(NG1,K),K=1,NTG1)
    4 FORMAT (24I3,8X)
      DO 10 I=1,NTG1
      IF(LEN.EQ.0) GO TO 25
      DO 20 J=1,LEN
      IF(MC(NG1,I).EQ.LIST(J)) GO TO 30
   20 CONTINUE
   25 LEN=LEN+1
      LIST(LEN)=MC(NG1,I)
      J=LEN
   30 MC(NG1,I)=J
   10 CONTINUE
    5 NTG(NG1)=NTG1
      NGM1=NG-1
C                                                                       C
C        THE NUMBER OF COMPARISONS WHICH MUST BE MADE BY                C
C        EACH JUDGE IS COMPUTED. NO COMPARISONS ARE MADE                C
C       BETWEEN OBJECTS IN THE SAME RANK.                               C
C                                                                       C
      DO 7 I=1,NGM1
      IP1=I+1
      NTGI=NTG(I)
      DO 7 J=IP1,NG
      NTGJ=NTG(J)
      DO 7 IC=1,NTGI
      DO 7 JC=1,NTGJ
      NN1=MC(I,IC)
      NN2=MC(J,JC)
      IF (NN1-NN2) 6,7,6
    6 WRITE(IDISK) NN1,NN2
      NCOUNT=NCOUNT+1
      NR(NN1)=1
      NC(NN2)=1
    7 CONTINUE
      NCNT1=NCOUNT-NCNT1
    8 WRITE(OUTPUT,9) J1,NCNT1
    9 FORMAT ( 6H0JUDGE I5, 10(1H ),I7,12H COMPARISONS )
      END FILE 9
      REWIND 9
```

```
      WRITE(OUTPUT,500)
  500 FORMAT(1H0,'ID NUMBER MAP')
      WRITE (OUTPUT,501)
  501 FORMAT(15(1H ),13HASSIGNED ID #,15(1H ),13HORIGINAL ID #/)
      WRITE(OUTPUT,502) (I,LIST(I),I=1,N)
  502 FORMAT(T20,I4,T49,I4)
      CALL CORE2
      CALL CORE3
      STOP
      END
C                                                                      C
C         RANKING PROCEDURE -- CORE2                                   C


      SUBROUTINE CORE2
      INTEGER*2 LIST(130)
      DIMENSION NR(130),NC(130),A(130,130),W(130),X(130),MAN(130)
      COMMON LIST
      COMMON M1,M2,INEW,JNEW,IP1,W,EPSLCN,NR,NC,N,NCOUNT,JUPPER,A,X,MAN
      INTEGER OUTPUT
      INPUT=5
      OUTPUT=6
      IDISK=9
      DO 1 I=1,N
      DO1 J=1,N
    1 A(I,J)=0
C                                                                      C
C         TABULATES FOR EACH COMPARISCN THE NUMBER OF TIMES            C
C         THAT COMPARISON IS MADE BY ALL JUDGES IN THE                 C
C         EXPERIMENT. THIS TABULATION IS DONE FROM RANKING             C
C         ORDER AND IS THE NUMBER OF TIMES OBJECT I WAS                C
C         PREFERRED TO OBJECT J,IE,THE WIN-LOSS MATRIX.                C
C                                                                      C
      DO 2 NCNT1=1,NCOUNT
      READ(IDISK) M1,M2
    2 A(M1,M2)=A(M1,M2)+1
      ITER=0
      DO 34 I=1,N
   34 MAN(I)=I
C                                                                      C
C         OBJECTS WHICH ARE RATED UNIVERSALLY HIGH OR                  C
C         UNIVERSALLY LOW ARE REMOVED FROM THE COMPUTATIONS.           C
C         THE APPROPRIATE ROWS AND COLUMNS OF THE WIN-LOSS             C
C         MATRIX ARE ALSO REMOVED.                                     C
C                                                                      C
   33 DO 3 I=1,N
      IF (NR(I)*NC(I)) 3,4,3
    3 CONTINUE
```

```
      GO TO 100
    4 INEW=0
      WRITE(OUTPUT,300)
  300 FORMAT(1H0,'WIN-LOSS MATRIX ORDERED BY ASSIGNED ID NUMBER')
      DO 70 I=1,N
   70 WRITE(OUTPUT,36) (A(I,J),J=1,N)
   36 FORMAT(30F4.0)
      ITER=ITER+1
      DC 10 I=1,N
      IF(NR(I)*NC(I)) 9,18,9
    9 INEW=INEW+1
      JNEW=1
      MAN(INEW)=MAN(I)
      DO 100C J=1,N
      IF (NR(J)*NC(J)) 11,1000,11
   11 A(INEW,JNEW)=A(I,J)
      JNEW=JNEW+1
 1000 CONTINUE
   10 CCNTINUE
      N=INEW
      DO 31 I=1,N
      NR(I)=0
   31 NC(I)=C
      DO 30 I=1,N
      DO 30 J=1,N
      IF (A(I,J)) 32,30,32
   32 NR(I)=1
      NC(J)=1
   30 CCNTINUE
      GO TO 33
C                                                                  C
C                                                                  C
C         COMPUTES THE INITIAL WEIGHTING FACTORS OF THE            C
C         OBJECTS FROM THE WIN-LOSS MATRIX. THESE WEIGHTING        C
C         FACTORS ARE THE PROBABILITY THAT OBJECT I IS
C         PREFERRED TO OBJECT J.                                   C
C                                                                  C
  100 DO 16 I=1,N
      W(I)=0
      DO 13 J=1,N
   13 W(I)=W(I)+A(I,J)
      Z=0
      DO 15 J=1,N
   15 Z=Z+A(J,I)
   16 X(I)=W(I)/(W(I)+Z)
      DC 200 I=1,N
      DO 200 J=1,N
  200 A(I,J)=A(I,J)+ .00001
```

```
C                                                                          C
C         COMPUTES THE NUMBER OF TIMES OBJECT I AND OBJECT J              C
C         ARE RANKED RELATIVE TO EACH OTHER.                              C
C                                                                          C
      DC 14 I=1,N
      IP1=I+1
      DO 14 J=IP1,N
      A(I,J)=A(I,J)+A(J,I)
   14 A(J,I)=A(I,J)
      REWIND 9
      RETURN
   18 IF (NR(I)) 8,5,8
    8 WRITE(OUTPUT,12) MAN(I),ITER
   12 FCRMAT(1HO,22HSUBJECT ASSIGNED ID #,I3,35H IS UNIVERSALLY RATED HI
     XGH,DELETED,I3, 3H RY )
      GC TO 10
    5 WRITE(CUTPUT,7) MAN(I),ITER
    7 FORMAT(1HO,22HSUBJECT ASSIGNED ID #,I3,35H IS UNIVERSALLY RATED LO
     XW ,DELETED,I3, 3H RY )
      GO TO 10
      END
C                                                                          C
C         RANKING PROCEDURE -- CORE3                                       C
C                                                                          C
      SUBROUTINE CORE3
      INTEGER*2 LIST(130)
      DIMENSION NR(130),NC(130),A(130,130),W(130),X(130),MAN(130),
     1 Y(130),Q(130)
      CCMMON LIST
      CCMMON JO,KC,DENOM,XA,XQ,W,EPSLCN,NR,NC,N,NCOUNT,JUPPER,A,X,MAN
      INTEGER OUTPUT
      INPUT=5
      CUTPUT=6
      IDISK=9
      JO=0
    8 KC=0
C                                                                          C
C         COMPUTES NEW WEIGHTING FACTOR FOR OBJECT I FOR                  C
C         ITERATICN BEING CONSIDERED.                                      C
C                                                                          C
      DO 3 I=1,N
      DENOM=C
      DO 4 J=1,N
    4 DENOM=DENOM+A(I,J)/(X(I)+X(J))
    3 Y(I)=W(I)/DENOM
      JO=JO+1
```

```
C
C          CONVERGENCE CRITERION CHECH.                                C
C                                                                      C
    5 DC 9 I=1,N                                                       C
      IF(ABS(Y(I)-X(I))/X(I)-EPSLON) 9,9,10
   10 KO=KO+1
    9 X(I)=Y(I)
      WRITE(OUTPUT,12) JO,KO,(Y(I),I=1,N)
   12 FCRMAT (1H0,I5,I10/(6F18.6))
      IF (KO) 17,17,13
C
C          ITERATICN COUNT CHECK.                                      C
C                                                                      C
   13 IF (JO-JUPPER) 8,14,14                                           C
   14 WRITE(OUTPUT,15) JUPPER
   15 FORMAT(1H0,I5,46H ITERATIONS, NO CONVERGENCE. DATA SET DELETED. )
   17 WRITE(CUTPUT,18) (MAN(I),LIST(MAN(I)),Y(I),I=1,N)
   18 FCRMAT (14H1FINAL WEIGHTS//15(1H ),
     X13HASSIGNED ID #, 3(1H ),13HORIGINAL ID #,12(1H ),6HWEIGHT//(1H0,
     X2I19,F24.6))
      RETURN
      END
```

BIBLIOGRAPHY

1. Abelson, R. M. and Bradley, R. A., "A 2x2 Factorial With Paired Comparisons," Biometrics, v. 10, p. 487-502, December 1954.

2. Battig, W. F. and Montague, W. E., "Category Norms for Verbal Items in 56 Categories," Journal of Experimental Psychology, v. 80, p. 1-46, June 1969.

3. Bradley, R. A. and Terry, M. E., "Rank Analysis of Incomplete Block Designs: The Method of Paired Comparisons," Biometrika, v. 39, p. 324-345, December 1952.

4. Bradley, R. A., "Some Statistical Methods in Taste Testing and Quality Evaluation," Biometrics, v. 9, p. 22-38, March 1953.

5. Bradley, R. A., "Incomplete Block Rank Analysis: On the Appropriateness of the Model for a Method of Paired Comparisons," Biometrics, v. 10, p. 375-390, September 1954.

6. Bradley, R. A., "Rank Analysis of Incomplete Block Designs: Some Large Sample Results on Estimation and Power for a Method of Paired Comparisons," Biometrika, v. 42, p. 450-470, December 1955.

7. Dixon, W. J., "Power Functions of the Sign Test and Power Efficiency for Normal Alternatives," Annals of Mathematical Statistics, v. 24, p. 467, 1953.

8. Ford, L. R., Jr., "Solution of a Ranking Problem from Comparisons," American Mathematical Monthly, v. 64, No. 8, Part II, p. 28-33, October 1957.

9. Guttman, L., "An Approach for Quantifying Paired Comparisons and Rank Order," Annals of Mathematical Statistics, v. 17, p. 144-163, June 1946.

10. Kendall, M. G. and Smith, B. B., "On the Method of Paired Comparisons," Biometrika, v. 31, p. 324-345, March 1940.

11. Ostle, B., Statistics in Research, 2d ed., p. 234, Iowa State University Press, 1963.

12. Pelz, D. C. and Andrews, F. M., Scientists in Organizations, Wiley, 1966.

13. Thorndike, E. L. and Lorge, I., The Teacher's Word Book of 30,000 Words, Columbia University Press, 1944.

14. White, C., "The Use of Ranks in a Test of Significance for Comparing Two Treatments," Biometrics, v. 8, p. 33-41, March 1952.